OIPE

```
This Raw Listing contains the General Information
Section and up to first 5 pages.
```

ENTERED

```
 1   <110> APPLICANT: Cohen, Daniel
 2         Blumenfeld, Marta
 3         Ilya, Chumakov
 4         Bougueleret, Lydie
 5   <120> TITLE OF INVENTION: PROSTATE CANCER GENE
 6   <130> FILE REFERENCE: GENSET.18CP1CP
 7   <140> CURRENT APPLICATION NUMBER: US/09/853,526
 8   <141> CURRENT FILING DATE: 2001-05-11
 9   <150> EARLIER APPLICATION NUMBER: 09/338,907
10   <151> EARLIER FILING DATE: 1999-06-23
11   <150> EARLIER APPLICATION NUMBER: 08/996,306
12   <151> EARLIER FILING DATE: 1997-12-22
13   <150> EARLIER APPLICATION NUMBER: 60/099,658
14   <151> EARLIER FILING DATE: 1998-09-09
15   <150> EARLIER APPLICATION NUMBER: 09/218,207
16   <151> EARLIER FILING DATE: 1998-12-22
17   <160> NUMBER OF SEQ ID NOS: 578
18   <170> SOFTWARE: Patent.pm
19   <210> SEQ ID NO 1
20   <211> LENGTH: 56516
21   <212> TYPE: DNA
22   <213> ORGANISM: Homo sapiens
23   <220> FEATURE:
24   <221> NAME/KEY: promoter
25   <222> LOCATION: 1629..1870
26   <223> OTHER INFORMATION: identification  method  Proscan
27   <220> FEATURE:
28   <221> NAME/KEY: misc_feature
29   <222> LOCATION: 1998..2000
30   <223> OTHER INFORMATION: potential start codon
31   <220> FEATURE:
32   <221> NAME/KEY: exon
33   <222> LOCATION: 2001..2216
34   <223> OTHER INFORMATION: exon1
35   <220> FEATURE:
36   <221> NAME/KEY: misc_feature
37   <222> LOCATION: 2031..2033
38   <223> OTHER INFORMATION: ATG
39   <220> FEATURE:
40   <221> NAME/KEY: misc_feature
41   <222> LOCATION: 11694..14332
42   <223> OTHER INFORMATION: Tyr  Phos
43   <220> FEATURE:
44   <221> NAME/KEY: primer_bind
```

```
45    <222> LOCATION: 11930..11947
46    <223> OTHER INFORMATION: upstream  amplification  primer  4-77  SEQ  ID42
47    <220> FEATURE:
48    <221> NAME/KEY: allele
49    <222> LOCATION: 12057..12103
50    <223> OTHER INFORMATION: polymorphic  fragment  4-77  SEQ  ID24
51    <220> FEATURE:
52    <221> NAME/KEY: primer_bind
53    <222> LOCATION: 12339..12358
54    <223> OTHER INFORMATION: downstream  amplification  primer  4-77  SEQ  ID51,  compl
55    <220> FEATURE:
56    <221> NAME/KEY: primer_bind
57    <222> LOCATION: 13547..13564
58    <223> OTHER INFORMATION: upstream  amplification  primer  4-73  SEQ  ID64
59    <220> FEATURE:
60    <221> NAME/KEY: allele
61    <222> LOCATION: 13657..13703
62    <223> OTHER INFORMATION: polymorphic  fragment  4-73  SEQ  ID58
63    <220> FEATURE:
64    <221> NAME/KEY: primer_bind
65    <222> LOCATION: 13962..13981
66    <223> OTHER INFORMATION: downstream  amplification  primer  4-73  SEQ  ID67,  compl
67    <220> FEATURE:
68    <221> NAME/KEY: exon
69    <222> LOCATION: 18196..18265
70    <223> OTHER INFORMATION: exon  2
71    <220> FEATURE:
72    <221> NAME/KEY: exon
73    <222> LOCATION: 23717..23832
74    <223> OTHER INFORMATION: exon  3
75    <220> FEATURE:
76    <221> NAME/KEY: exon
77    <222> LOCATION: 25571..25660
78    <223> OTHER INFORMATION: exon  4
79    <220> FEATURE:
80    <221> NAME/KEY: primer_bind
81    <222> LOCATION: 34216..34234
82    <223> OTHER INFORMATION: upstream  amplification  primer  99-217  SEQ  ID43
83    <220> FEATURE:
84    <221> NAME/KEY: allele
85    <222> LOCATION: 34469..34515
86    <223> OTHER INFORMATION: polymorphic  fragment  99-217  SEQ  ID25
87    <220> FEATURE:
88    <221> NAME/KEY: primer_bind
89    <222> LOCATION: 34625..34645
90    <223> OTHER INFORMATION: downstream  amplification  primer  99-217  SEQ  ID52,  com
91    <220> FEATURE:
92    <221> NAME/KEY: exon
93    <222> LOCATION: 34669..34759
94    <223> OTHER INFORMATION: exon  5
```

```
 95   <220> FEATURE:
 96   <221> NAME/KEY: exon
 97   <222> LOCATION: 40688..40846
 98   <223> OTHER INFORMATION: exon  6
 99   <220> FEATURE:
100   <221> NAME/KEY: exon
101   <222> LOCATION: 48070..48193
102   <223> OTHER INFORMATION: exon  7
103   <220> FEATURE:
104   <221> NAME/KEY: exon
105   <222> LOCATION: 50182..54523
106   <223> OTHER INFORMATION: exon  8
107   <220> FEATURE:
108   <221> NAME/KEY: primer_bind
109   <222> LOCATION: 51149..51168
110   <223> OTHER INFORMATION: upstream  amplification  primer  4-65  SEQ  ID65
111   <220> FEATURE:
112   <221> NAME/KEY: allele
113   <222> LOCATION: 51448..51494
114   <223> OTHER INFORMATION: polymorphic  fragment  4-65  SEQ  ID59
115   <220> FEATURE:
116   <221> NAME/KEY: primer_bind
117   <222> LOCATION: 51482..51499
118   <223> OTHER INFORMATION: downstream  amplification  primer  4-65  SEQ  ID68,  compl
119   <220> FEATURE:
120   <221> NAME/KEY: primer_bind
121   <222> LOCATION: 51596..51613
122   <223> OTHER INFORMATION: upstream  amplification  primer  4-67  SEQ  ID44
123   <220> FEATURE:
124   <221> NAME/KEY: allele
125   <222> LOCATION: 51612..51658
126   <223> OTHER INFORMATION: polymorphic  fragment  4-67  SEQ  ID26
127   <220> FEATURE:
128   <221> NAME/KEY: primer_bind
129   <222> LOCATION: 51996..52015
130   <223> OTHER INFORMATION: downstream  amplification  primer  4-67  SEQ  ID53,  compl
131   <220> FEATURE:
132   <221> NAME/KEY: polyA_signal
133   <222> LOCATION: 54445..54450
134   <223> OTHER INFORMATION: AATAAA
135   <400> SEQUENCE: 1
136         gtggatctgt gactgttcgc aggaagagag gagcgggagc aggacagaca ataactgata    60
137         gtcaggagct gggtttggag ataaagaggg aacaagagaa agttaagttc tgtgttttca   120
138         tggcaaacat tgcacaaaag tttacaactt cgtgactaac agtaatctgg ggtgattcac   180
139         aacaaattta cacataaaca catatttact gactttatac acagcaatcc taacgtgaac   240
140         acagaacctg ctttatcttt cgcacactg ttctagtgta gagatgtctg tctcagtta    300
141         aagaaagcat aaggagcatt agttgtgcac actgtccaca cccgtgactt ttttccacca   360
142         gtactaaacc tagtgcttct tacagtacag ggcaatgaca gccacagaaa gagagaagct   420
143         cctttactg tgtaatgctt cctgctggcc ttcaaatact tgttacttga gagatctcca   480
144         ttcacctggc tttgtcccca aaggtcatca tctaccaatg atgttgttat ttgatgttaa   540
```

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/853,526

DATE: 01/17/2002
TIME: 12:43:35

OIPE

Input Set: **I853526.RAW**

```
145     tcatgtataa agaaagtagc taccatcctg gccctgatta gaacttccca ctgaaatacc      600
146     gtcctgccta aaggtagcac aggtttccat tatggtggtg gtggggaggg ggcgggaata      660
147     tatatatata tatatatata tatatatatg gtaaagcatt cggcattctt ttaaagtaca      720
148     actatccttg aaaagggtta catattaaac cattttacc acagccaaag gggaggagaa        780
149     agatccaaaa gtcctgtgga tctgctttaa catcaataaa acagttatcc acccttcgta      840
150     gcttttagtg aaggctacaa aagtatgctt tttatggatt acacatgtgc acgcaactac      900
151     tttaattact acagaaaaaa acgaggctcc ttattaaaaa aaaatcagaa acaagtccaa      960
152     cagactctga ggaaatgaag caagagtgaa ttctgaaaag gtctaataaa cagtatggaa     1020
153     atatccttgt gggattgttc ttcagctatg cataaacatg taattatcat cattactgtg     1080
154     atggggaaaa acacggaccc taattctgaa acaccctggt agcgagagac gggcaggagg     1140
155     ggctgctgcg cactcagagc ggaggctgag gaggcggcgt ccccttgcaa aggactggca     1200
156     gtgagcagat ggggacactc gagctgcccc gcgacctggg ccgagctgcc tacaacctgg     1260
157     gcccaggtgc ctgcaagaat tagacctccg ataacgttaa cacccacttt ctcactgctc     1320
158     taattgtgtg catcccggcg cccaggggct tgtgagcagc aggtgcgcgt tccaggcagc     1380
159     tccagcgacc cttaaacctg accgcgcgca cgtccggccc gagggagcag aacaagaggc     1440
160     acccggaccc tcctccggcc agcacccacc ttcacccagt tccgtcagtc gccaccacct     1500
161     cccttcccgc gtccgcagcc ggcccagctg gggagcatgc gcagtggccg gagccgggtt     1560
162     gcccgcgcca cagcaggtag ctgtactgca actgtcggcc caaaccaacc aatcaagaga     1620
163     cgtgttattg ccgccgaggt ggaactatgg caacgggcga ccaatcagaa ggcgcgttgt     1680
164     tgccgcggag ccccctgccc cggcagggg atgtggcgat gggtgagggt catggggtgt      1740
165     gagcatccct gagccatcga tccgggaggg ccgcgggttc ccttgctttg ccgccgggag     1800
166     cggcgcacgc agccccgcac tcgcctaccc ggccccgggc ggcggcgcgg cccatgcggc     1860
167     tgggggcgga ggctgggagc gggtggcggg cgcggcggcc cgggcccggg cggtgattgg     1920
168     ccgcctgctg gccgcgactg aggcccggga ggcgggcggg gagcgcaggc ggagctcgct     1980
169     gccgccgagc tgagaagatg ctgctgtccc tggtgctcca cacgtactcc atgcgctacc     2040
170     tgctgcccag cgtcgtgctc ctgggcacgg cgcccaccta cgtgttggcc tggggggtct     2100
171     ggcggctgct ctccgccttc ctgcccgccc gcttctacca agcgctggac gaccggctgt     2160
172     actgcgtcta ccagagcatg gtgctcttct cttcgagaa ttacaccggg gtccaggtga      2220
173     gccgcctccc gctcccgggt ctcggcgtcc acccgagctc ccggggggcgc ggacctctcc     2280
174     gctcccccac agctggcgag ggtcacccgg ccggcccggc ggaccagca cggagagcac      2340
175     gtgccgcctc cccgccttcc tctccgcatg cttcctgccg ttctgccgag atcgctctct     2400
176     aggaagctgt ggctgcgtcg tcctgaggct acgagtggga cccgccgccc ctttccccgc     2460
177     ccctcgcctg ggtctgatgc tgcttagcaa agtgggtgca gatgcacgtt ttaaataata     2520
178     gggcacgcgt ttagcagttt ctggcctttg gtccaaagag gtggtcatgt tggaacagat     2580
179     cggagacgtc tacactccga agtgcgcttt tacagtgacc tcttgaaaca gaagtacaat     2640
180     tcggtcttgt gttctttccc ctggacaagt gaaagctggg cgaagaaatg aatacatttg     2700
181     ttaaccgtag aagcctaact agatacaatt cttgccaact ttaactgggc ttgaatgtgt     2760
182     gggtgatctg ttgtctgatt actttctttc tgttactgtt tctctgtaga gattggattc     2820
183     gtagattaaa cttgagaaac aaaccataaa agtggaaggc cctctttaac agtaggtatt     2880
184     tgaagtgtta taaaaaaaaa aaaggtgaat ttttctttta tttctcagtt tgaaagaaca     2940
185     gctttattct tggttattcc taatgtccac ctagtcctct tttacttttc ttggtagggt     3000
186     tagggtggca tggggaaatg ggacggtatc attttgtctt tttaacttttt ttttttttcca     3060
187     cctacagcag ctgtttttac cctgtggtca gtcaggtact atatttagtt tgcagttgca     3120
188     ctgctgatcg acccttgatg gccccagttg gaagttgttt gggggggaagg aactaggaga     3180
189     ggccagggcc tccatttaaa ccagtgtctg taagtgtctc cttggaagga aaaaagata      3240
190     ctgttccagg tcatggtttc ctggtagttg acgtttaaaa tgggcctcat ttaaaaattt     3300
191     caataattca ggctaatttt ttccctttat atggtaactc caccaagttt gtctaaatgt     3360
192     atgatttta tcatgattaa gtttttactt ccacatcatg tgcaactgg cctgggatgg      3420
193     gatataagct cagaacacaa agtcattcac ctgttaaaaa aataattcta tctgtggcgg     3480
194     gttatgttat ttttgttcaa agaggacaca atatgatgca gaatacacca ttgaaggatt     3540
```

```
195     ttttggtttg gcaagttctt attttttaa atggctgtaa aacctagcag tgtttctgaa      3600
196     attgcatacc ttacctgatg ttcagagatc cgatttactt cttgatttcc cagcaagtga      3660
197     ttttgaaaac atttaatcta atcattcccc ccaccgtctg ttcaaatcaa aggaagtggc      3720
198     atccagcact aattttcatg catttatgaa aggatgcctg aggacccta agtataattc       3780
199     aaaattttgt ttaatgtgtg ttccttgatg aagttcttta ggagtcgtag aacgaactga      3840
200     ttgcccactg atcatcaaat gcaagttatg aacatttaat aaaaatttaa aaccaagagt      3900
201     ttcttgttcc tgcattttta ttttattgt atggagggga caaataatta ttttctgttt       3960
202     agtaacagag cagggtattt tgaatttatt agggtctttt tctgcagtct gggtttcctg      4020
203     tgtacacaaa gctacctttc aatattttt attgtttctg ttaagattaa atcaatagag       4080
204     gaataaatag ctatcttcaa acataagacc caaaggaaaa agatttatag tgatgttctg      4140
205     tcaccttatt ttttacctgt gactttgtac cattaacttt gtcactgaga tgttttgatt      4200
206     aaaatttta gcttgctttt cttgttttgt taggacactc ttttttttt gaattgtttt        4260
207     tatcagcttt cgtttgcaag gctagtgatg attctcttgt tctgtataaa gtattgttga      4320
208     ctcatttctg aagggagttt tagtaattta agaggttata agttttaaa taaaaggttt       4380
209     attaatttat atatattaaa gaggcatttt aaaataaat ttttttaaa tgacattttt        4440
210     acacctttca actctaggtt taaaaaataa gtggttcaca gtagttcttg cagaagaata      4500
211     ttttctttta catagaattt ttaagctgaa gagaagtagt agtaggtcca tgagatttat      4560
212     gatctgtgct tggcaggtaa acctgcttcc aacaaattta gttggatttt cttggattc       4620
213     tgggtaaata ccttttctt ccccagtttc actactttat tttcatatgt atctctgaga       4680
214     tagagaaata tttcagtcag tgctgctaaa attgttcctt ataactcgtt tatccttta       4740
215     ggtccttcca gaatctctca ttggtactga aactcaaatg ggtactttct tcaccattta      4800
216     tttctttaga ataagtaata agaatttat aagctttttt atatttcacg taatttgaga       4860
217     ctattgaaaa tccagttaag tctctctact gtgttgagag gcattgattc aagtacctgt      4920
218     gttacttttcc tgtgctgcca aaacagatca cctcaaacta agcggcttaa aataatagaa     4980
219     cttaagttct cgtgattctg gaggccagca ctttgaaatc aaggtgtagg ctcaatttta      5040
220     ctccctctgg aggccctagg gggaatctgt tcttgtgggt ttcaacttct ggtgactggt      5100
221     ggcattcctt ggcttggggc cccatcactt caacctctgc cttacagtcc ttgctgccac      5160
222     ctcttctgtc tcacatctca ctctcccttt ctcttagaag gatgcttgtc attgggttta      5220
223     gagcccacct ggatattccg ggatgatctc ttcatctcaa gatccttaat tataactgca      5280
224     aagagccttt ttccaaataa gaaaacattc acaggttcca gggcttagga tgtggacaca      5340
225     tttttgagg ggctgcccctt cattcccca caacaatgaa ctccatagtt ctgcctattc       5400
226     agtattttgt agttatttcg tagtttaact tgccttattt ctttaggtat ttacgtatta      5460
227     aagcattttg gtctctgctt tctttaacag agaacctggt tttctgtaat aagtttactt      5520
228     actttcccat aatctttag tttcttattt acagatttac cttcacatat cccttaagta       5580
229     gaacatttga ttaactgttt tattttcgga acaaatctgc attctgtata ataaccaact      5640
230     tattcatatt tcggtattct tttaattctt atctgattct gaaattacca tcttgtgatt      5700
231     atatatatat atatatggaa ataactgaaa tcttgataaa ttaaaggtga tataacttct      5760
232     aagacaatta attatgtatg atgtggtgaa tatactggtg tttggtttgt ttgccactta      5820
233     aaagccctat ctataggata ggaagtaact tgaatgtgga atgcttagag actcagagta      5880
234     agaggccgta tatatatcct tgagctggag tttaaggaaa acttatggga aattaaaagg      5940
235     aaagttggag tactgacaga ggattgcgta ggactcatga aaaaggaatg aagttacctt      6000
236     aaattctatc atcgtgagtt aacgtgaaac tagatttatg ttagtttata gcctagaatt      6060
237     ctatcctagg aatctagata tatcctaaat gttgagatag ctgcataaac aataactgta      6120
238     atcgttatga taaataatga caaatctttt tagcatgttt tgtgaagctg ataaatgtta      6180
239     ataggatgtc ttcaaatgtc agaattcttt tttctttgct tctttttaa aaaatttctt       6240
240     ttcccccatt cctatgcaat acactgaaaa ctgatcattg aaatttgtag gccaaaaaat      6300
241     taatcaacac gtaatagatt ggggtttggg tttttttgag tcaggtcttc ttctgtcac       6360
242     ccaggctctg gtgcggtggc accatcatgg ctcattgcag ccttgaatgc ctgggttcaa      6420
243     gtgatcctcc ggagtagctg ccgtgccatt atttctagct aatttttaaa agttttgta       6480
244     gaaatggggt ctttctgtgt tgcccaggct ggtcttgaat cctggcctc aggtgatcct        6540
```

VERIFICATION SUMMARY
PATENT APPLICATION US/09/853,526

Input Set: I853526.RAW

Line	?	Error/Warning	Original Text
435	W	"N" or "Xaa" used: Feature required	atgcttgcta tattttctc atgaggatat tgattatt
436	W	"N" or "Xaa" used: Feature required	nnntgtacta tacataactg ctttctgtac ctgagcta
458	W	"N" or "Xaa" used: Feature required	attacattca gatttatagg acaagggttg aagctanc
787	W	"N" or "Xaa" used: Feature required	cctctctctc ttttctttct gcnnnnnnnn nnnattct
1160	W	Line data has been corrected	Gly Cys Leu Trp Val Thr Ile Lys Ala *
1230	W	"N" or "Xaa" used: Feature required	atatgcattt caaaaanaaa aaaaaaaaa aaaaa
1646	W	"N" or "Xaa" used: Feature required	ccctgtnaga cacgtcctgt atcgttgttg agatggga
1844	W	"N" or "Xaa" used: Feature required	ccctgtnaga cacgtcctgt atcattgttg agatggga
2425	W	Line data has been corrected	Val Thr Val Val Tyr Glu Gly Lys Asp Asp G
2501	W	"N" or "Xaa" used: Feature required	tcattaattc agcaataaag gaaaatatgc atttcaaa
2625	W	Line data has been corrected	Tyr Gly Thr Leu Leu Gly Cys Leu Trp Phe V
2626	W	Invalid/Missing Amino Acid Numbering	345 350
2629	W	"N" or "Xaa" used: Feature required	taagccttga tgatngnncn cnnnnnnnnn ncnantcn
2678	W	"N" or "Xaa" used: Feature required	tgacacacct gccttcatct ccaaggtaca ggaattgc
2679	W	"N" or "Xaa" used: Feature required	nntttttttt tttgagtttt ggggaggggg tatatttt
2699	W	"N" or "Xaa" used: Feature required	ccccaaggtc aggatgtggg tggtactgct gtctccca
2752	W	"N" or "Xaa" used: Feature required	taggtgtnct cacaaatacc ccaagctaaa caaccaca
2766	W	"N" or "Xaa" used: Feature required	agatagnttg gaatcttggc tctcatgtag ttgtactg
2767	W	"N" or "Xaa" used: Feature required	gacctttntg gaaagaacat ataactggga ggggcttt
2768	W	"N" or "Xaa" used: Feature required	ttcctagttc aaactntact gcctgctcaa agcttgag
2770	W	"N" or "Xaa" used: Feature required	gaattttaag tataaataaa cncttctttn taaaacaa
2799	W	"N" or "Xaa" used: Feature required	gactattagc atattctaag aaacagtttt naggattt
2809	W	"N" or "Xaa" used: Feature required	atccccaaga aagtatactt ttnggagtga ctcagtat
2810	W	"N" or "Xaa" used: Feature required	acacgcacgt gcggatgtat atgtagcaca tatgtaaa
2824	W	"N" or "Xaa" used: Feature required	tctgtaaagg tcctgagaca agatggcctc tagtggcg
2842	W	"N" or "Xaa" used: Feature required	cccagcacca atcacccact cttatnctttt tatgcntt
2877	W	"N" or "Xaa" used: Feature required	caatgtaaga aggcctatgt atcggtaact cactgttg
2893	W	"N" or "Xaa" used: Feature required	ggtatacaga cacacacact agccaacacc catctaca
2906	W	"N" or "Xaa" used: Feature required	ggggttgctg gagacatgga agctaagtgg agtttatg
3453	W	Line data has been corrected	Ser Gly Arg Met Arg *
3563	W	Line data has been corrected	Ser Gly Arg Met Arg *
3705	W	Line data has been corrected	Trp Val Thr Ile Lys Ala *
3842	W	Line data has been corrected	Leu Leu Gly Cys Leu Trp Val Thr Ile Lys A
3973	W	Line data has been corrected	Gly Thr Leu Leu Gly Cys Leu Trp Val Thr I
4074	W	Line data has been corrected	Asn Ile *
4173	W	Line data has been corrected	Asn Tyr Thr Gly Val Gln Asp Ala Tyr Arg I
4271	W	Line data has been corrected	Asn Tyr Thr Gly Val Gln Arg Leu Asp Ser
4387	W	Line data has been corrected	Cys Asp Phe Ser Arg Arg Tyr Lys Val *
4512	W	Line data has been corrected	Ile Val Ser Thr Lys Lys Met Ser Gln Lys A
4659	W	Line data has been corrected	Ile Lys Ala *
4798	W	Line data has been corrected	Thr Ile Lys Ala *
4916	W	Line data has been corrected	Gly *
7266	W	"N" or "Xaa" used: Feature required	cctgctactg tttctttctt tctttccanc cctccctc
7290	W	"N" or "Xaa" used: Feature required	tggagagcag tgttcaagca gctgtatcca accaattn
7335	W	"N" or "Xaa" used: Feature required	ggaaaaaaaa aaaaanaacc cttttgaaat aaggacag
7350	W	"N" or "Xaa" used: Feature required	cacacactct ttttccctgt atgccttcac tcctcgaa
7371	W	"N" or "Xaa" used: Feature required	gagagggagg gagggagaga gagagagaga gagagaga
7372	W	"N" or "Xaa" used: Feature required	nnncntnna nnncntggaa caaattaaaa actatgta
7429	W	"N" or "Xaa" used: Feature required	tctagcnttt tatttactct gagacatatt cttgctct
7469	W	"N" or "Xaa" used: Feature required	tgaatgatat ctattgatag aagcttnang nnnnnnnn

VERIFICATION SUMMARY
PATENT APPLICATION **US/09/853,526**

Input Set: **I853526.RAW**

Line	?	Error/Warning	Original Text
7476	W	"N" or "Xaa" used: Feature required	gnacagtatt gaatctatat tgtattcttt gactgtac
7529	W	"N" or "Xaa" used: Feature required	gccagcttgt ccacatagtt ccaggagagc cagggcta
7617	W	"N" or "Xaa" used: Feature required	catcacaaga ccatgtgcgg tagtttggaa aacagtct
7618	W	"N" or "Xaa" used: Feature required	ggaatcttgg ctctcatgta gttgtactga ttagatca
7620	W	"N" or "Xaa" used: Feature required	caaactntac tgcctgctca aagcttgagg catgaact
7622	W	"N" or "Xaa" used: Feature required	gataaataaa cacttctttt aaaacaacaa gaacaaca
7957	W	"N" or "Xaa" used: Feature required	ttcctgtctc naaaaaaaaa aaaaaaaaa
8033	W	Line data has been corrected	Tyr Gly Thr Leu Leu Gly Cys Leu Trp Phe V
8034	W	Invalid/Missing Amino Acid Numbering	345 350